

(12) 专利合作条约所公布的国际专利

(19) 世界知识产权组织  
国际局

(43) 国际公布日:

2003年7月17日(17.07.2003)

PCT

(10) 国际公布号:

WO 03/057883 A1

- (51) 国际分类号: C12N 15/12
- (21) 国际申请号: PCT/CN03/00003
- (22) 国际申请日: 2003年1月2日(02.01.2003)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权: 02110535.9 2002年1月11日(11.01.2002) CN
- (71)(72) 发明人/申请人: 余龙(YU, Long) [CN/CN]; 中国上海市邯郸路220号复旦大学遗传学研究所, Shanghai 200433 (CN).
- (72) 发明人; 及
- (75) 发明人/申请人(仅对美国): 唐丽莎(TANG, Lisha) [CN/CN]; 郭泽坤(GUO, Zekun) [CN/CN]; 章平肇(ZHANG, Pingzhao) [CN/CN]; 董轶敏(DONG, Yimin) [CN/CN]; 中国上海市邯郸路 220号复旦大学遗传学研究所, Shanghai 200433 (CN).
- (74) 代理人: 上海专利商标事务所(SHANGHAI PATENT & TRADEMARK LAW OFFICE); 中国上海市桂平路435号徐迅, Shanghai 200233 (CN).
- (81) 指定国(国家): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW
- (84) 指定国(地区): ARIPO专利(GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), 欧亚专利(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), 欧洲专利(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI专利(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)
- 本国际公布:  
— 包括国际检索报告。
- 所引用双字母代码和其它缩写符号, 请参考刊登在每期PCT公报期刊起始的“代码及缩写符号简要说明”。

(54) Title: THE HUMAN HEPATOMA-DERIVED GROWTH FACTOR 5, ITS CODING SEQUENCE, THE METHOD OF PRODUCING IT AND THE USES OF IT

(54) 发明名称: 人类肝癌衍生生长因子5、其编码序列、制法及用途

(57) Abstract: The present invention provides the cDNA sequence of human hepatoma-derived growth factor (HDGF5). The invention also relates to the polypeptide encoded by the polynucleotide, to the uses thereof, to the method for producing them.

(57) 摘要

本发明提供了一种人肝癌细胞衍生生长因子(HDGF5)的cDNA序列。本发明还涉及由该核苷酸序列编码的多肽、这些多核苷酸和多肽的应用以及所述多核苷酸和所述多肽的制备方法。

## 人类肝癌衍生生长因子 5、其编码序列、制法及用途

### 技术领域

5 本发明涉及基因工程领域，具体地，涉及一种人基因核苷酸序列。更具体地说，本发明涉及一种人肝癌细胞衍生生长因子(HDGF5)的 cDNA 序列。本发明还涉及由该核苷酸序列编码的多肽、这些多核苷酸和多肽的应用以及所述多核苷酸和所述多肽的制备方法。

### 背景技术

10 生长因子是调节细胞增殖和分化的一类物质，其分子量从几百到几万不等，习惯上称多肽生长因子，通过与特异的膜表面受体作用后引发的一系列级联反应而实现对细胞生长的调节。与经典的多肽和蛋白质激素相比，多肽生长因子没有特定的内分泌腺和内分泌细胞，而是通过一些细胞的旁分泌和自分泌释放并扩散到靶细胞，从而协调机体自身的统一和对外界的反应。

15 肝癌衍生生长因子(hepatoma derived growth factor, HDGF)是一种较为重要的生长因子，这个名称首次出现 Klagsbrun,M.等人 1986 年的论文 (P.N.A.S. USA Vol.33, pp2448-2452, 1986)中。在该文中，Klagsbrun,M.等人从人的肝癌细胞株 SK-HEP-1 分离纯化了一个约 18.5-19KD 蛋白因子，这种因子的一个显著特点就是与肝素有很强的亲和力。

20 在 1989 年，日本大阪大学医学院的 Nakamura 实验小组成员从人的肝癌细胞株 HuH-7 分离纯化到一个约 64KD 的蛋白因子，也命名为 HDGF(Clinic Chimica Acta.,183 : 273-284 , 1989)。在该文中，Nakamura,H.等人对该因子的生化性质和功能作了初步的研究，并认为 HDGF 不同于当时已发现的 PDGF、FGF、HGF(肝癌生长因子)等细胞因子。1997 年该实验组在小鼠中找到了 HDGF 的同系物并发现了该基因家族的另两个成员 HRP-1、HRP-2，它们都具有一个相当保守的 98 个氨基酸长的 HATH (homologous to the amino terminus of HDGF)序列 (Biochem. Biophys. Res. Commun. 238:26-32 , 1997)，1999 年该实验组又在人和小鼠中克隆到了 HDGF 基因家族的另一个成员 HRP-3(Biochem. Biophys. Res. Commun. 266(1):81-87 , 1999)。

30 然而，在本发明被公布之前，尚没有任何人公开过本申请中涉及的人 HDGF5。

### 发明内容

本发明的目的在于提供一种新的人肝癌细胞衍生生长因子多核苷酸序列，该

多核苷酸序列被命名为人 HDGF5。

本发明的另一个目的是提供一种新的人肝癌细胞衍生生长因子的蛋白，该蛋白被命名为人 HDGF5。

5 本发明的再一个目的是提供一种利用重组技术生产所述的新的入 HDGF5 蛋白的方法。

本发明还提供了这种人 HDGF5 基因序列和蛋白的应用。

10 在本发明的一个方面，提供了一种分离出的 DNA 分子，它包括：编码具有人 HDGF5 蛋白活性的多肽的核苷酸序列，所述的核苷酸序列与 SEQ ID NO:1 中从核苷酸 5-910 位的核苷酸序列有至少 70%的同源性；或者所述的核苷酸序列能在中度严谨条件下与 SEQ ID NO:1 中从核苷酸 5-910 位的核苷酸序列杂交。较佳地，所述的序列编码一多肽，该多肽具有 SEQ ID NO:2 所示的序列。更佳地，该序列具有 SEQ ID NO:1 中从核苷酸 5-910 位的核苷酸序列。

15 在本发明的另一方面，提供了一种分离的 HDGF5 蛋白多肽，它包括：具有 SEQ ID NO:2 氨基酸序列的多肽、或其保守性变异多肽、或其活性片段、或其活性衍生物。

较佳地，该多肽选自下组：(a)具有 SEQ ID NO: 2 氨基酸序列的多肽；(b)将 SEQ ID NO:2 氨基酸序列经过一个或多个氨基酸残基的取代、缺失或添加而形成的，且具有促进细胞增殖功能的由(a)衍生的多肽。

20 在本发明的另一方面，提供了一种载体，它含有上述分离出的 DNA。

在本发明的另一方面，提供了一种所述载体转化的宿主细胞。

在本发明的另一方面，提供了一种产生具有 HDGF5 蛋白活性的多肽的方法，该方法包括：

25 (a)将编码具有 HDGF5 蛋白活性的多肽的核苷酸序列可操作地连于表达调控序列，形成 HDGF5 蛋白表达载体，所述的核苷酸序列与 SEQ ID NO:1 中从核苷酸 5-910 位的核苷酸序列有至少 70%的同源性；

(b)将步骤(a)中的表达载体转入宿主细胞，形成 HDGF5 蛋白的重组细胞；

(c)在适合表达 HDGF5 蛋白多肽的条件下，培养步骤(b)中的重组细胞；

(d)分离出具有 HDGF5 蛋白活性的多肽。

30 在本发明的另一方面，提供了一种抗体，它与HDGF5蛋白多肽特异性结合。

在本发明的另一方面，提供了一种探针分子，它含有如SEQ ID NO:1所示的DNA分子中8-100个连续核苷酸。

附图说明

图1是人HDGF5与HDGF家族其他成员氨基酸N端序列同源比较。其中，“\*”代表各序列在该位置具有相同的氨基酸，“:”表示各序列在该位置的氨基酸分布具有较高的同源性，“.”表示各序列在该位置的氨基酸分布具有较高的同源性。附图中各序列的Genbank注册号如下：

大鼠HRP	AAL29938
小鼠HRP1	JC5661
小鼠HDGF	JC5660
大鼠HDGF	AAL47132
人HDGF	D16431
人HRP1	AAH18991
人HRP3	BAA90477
小鼠HRP3	AB029493
小鼠HRP2	JC5662
大鼠HDGF3	AAK50635

5

由图1可知，本发明的人HDGF5其氨基末端(HATH区域通常包括N端98个氨基酸)与HDGF家族其他成员都有较高同源性。

### 具体实施方式

10 在本发明中，“分离的”、“纯化的”或“基本纯的”DNA是指，该DNA或片段已从天然状态下位于其两侧的序列中分离出来，还指该DNA或片段已经与天然状态下伴随核酸的组份分开，而且已经与在细胞中伴随其的蛋白质分开。

在本发明中，术语“HDGF5蛋白(或多肽)编码序列”指编码具有HDGF5蛋白活性的多肽的核苷酸序列，如SEQ ID NO:1中5-910位核苷酸序列及其简并序列。该简并序列是指，位于SEQ ID NO:1序列的编码框5-910位核苷酸中，有一个或多个密码子被编码相同氨基酸的简并密码子所取代后而产生的序列。由于密码子的简并性，所以与SEQ ID NO:1中5-910位核苷酸序列同源性低至约70%的简并序列也能编码出SEQ ID NO:2所述的序列。该术语还包括能在中度严紧条件下，更佳地，在高度严紧条件下与SEQ ID NO:1中从核苷酸5-910位的核苷酸序列杂交的核苷酸序列。此外，该术语还包括与SEQ ID NO:1中从核苷酸5-910位的核苷酸序列的同源性至少70%，更佳地至少80%，更佳地至少90%的核苷酸序列。该术语还包括能编码具有与人HDGF5相同功能的蛋白的SEQ ID NO:1中开放读框序列的变异形式。这些变异形式包括(但不限于)：若干个(通常为1-90

个, 较佳地 1-60 个, 更佳地 1-20 个, 最佳地 1-10 个)核苷酸的缺失、插入和/或取代, 以及在 5'和/或 3'端添加数个(通常为 60 个以内, 较佳地为 30 个以内, 更佳地为 10 个以内, 最佳地为 5 个以内)核苷酸。本发明的编码序列可以是 DNA 或 RNA, 可以是单链或双链。

5 在本发明中, “基本纯的”蛋白质或多肽是指其至少占样品总物质的 20%, 较佳地 50%, 更佳地 80%, 最佳地 90%(按干重或湿重计)。纯度可以用任何合适的方法进行测量, 如用柱层析、PAGE 或 HPLC 法测量多肽的纯度。基本纯的多肽基本上不含天然状态下的伴随其的组分。

10 在本发明中, 术语“HDGF5 蛋白多肽”指具有 HDGF5 蛋白活性的 SEQ ID NO:2 序列的多肽。该术语还包括具有与人 HDGF5 相同功能的、SEQ ID NO:2 序列的变异形式。这些变异形式包括(但并不限于): 若干个(通常为 1-50 个, 较佳地 1-30 个, 更佳地 1-20 个, 最佳地 1-10 个)氨基酸的缺失、插入和/或取代, 以及在 C 末端和/或 N 末端添加一个或数个(通常为 20 个以内, 较佳地为 10 个以内, 更佳地为 5 个以内)氨基酸。例如, 在本领域中, 用性能相近或相似的氨基酸进行取  
15 代时(最好按表 1 所示进行), 通常不会改变蛋白质的功能。又比如, 在 C 末端和/或 N 末端添加一个或数个氨基酸通常也不会改变蛋白质的功能。该术语还包括 HDGF5 蛋白的活性片段和活性衍生物。

20 该多肽的变异形式包括: 同源序列、保守性变异体、等位变异体、天然突变体、诱导突变体、在高或低的严紧度条件下能与 HDGF5 DNA 杂交的 DNA 所编码的蛋白、以及利用抗 HDGF5 多肽的抗血清获得的多肽或蛋白。本发明还提供了其他多肽, 如包含 HDGF5 多肽或其片段的融合蛋白。除了几乎全长的多肽外, 本发明还包括了 HDGF5 多肽的可溶性片段。通常, 该片段具有 HDGF5 多肽序列的至少约 10 个连续氨基酸, 通常至少约 30 个连续氨基酸, 较佳地至少约 50 个连续氨基酸, 更佳地至少约 80 个连续氨基酸, 最佳地至少约 100 个连续氨基酸。

25 发明还提供 HDGF5 蛋白或多肽的类似物。这些类似物与天然 HDGF5 多肽的差别可以是氨基酸序列上的差异, 也可以是不影响序列的修饰形式上的差异, 或者兼而有之。这些多肽包括天然或诱导的遗传变异体。诱导变异体可以通过各种技术得到, 如通过辐射或暴露于诱变剂而产生随机诱变, 还可通过定点诱变法或其他已知分子生物学的技术。类似物还包括具有不同于天然 L-氨基酸的残基(如  
30 D-氨基酸)的类似物, 以及具有非天然存在的或合成的氨基酸(如  $\beta$ 、 $\gamma$ -氨基酸)的类似物。应理解, 本发明的多肽并不限于上述列举的代表性的多肽。

修饰(通常不改变一级结构)形式包括: 体内或体外的多肽的化学衍生形式如乙酰化或羧基化。修饰还包括糖基化, 如那些在多肽的合成和加工中或进一步加工步骤中进行糖基化修饰而产生的多肽。这种修饰可以通过将多肽暴露于进行糖

基化的酶(如哺乳动物的糖基化酶或去糖基化酶)而完成。修饰形式还包括具有磷酸化氨基酸残基(如磷酸酪氨酸, 磷酸丝氨酸, 磷酸苏氨酸)的序列。还包括被修饰从而提高了其抗蛋白水解性能或优化了溶解性能的多肽。

在本发明中, “人HDGF5保守性变异多肽”指与SEQ ID NO: 2的氨基酸序列相比, 有至多10个, 较佳地至多8个, 更佳地至多5个, 最佳地至多3个氨基酸被性质相似或相近的氨基酸所替换而形成多肽。这些保守性变异多肽最好根据表1进行氨基酸替换而产生。

表 1

最初的残基	代表性的取代	优选的取代
Ala (A)	Val; Leu; Ile	Val
Arg (R)	Lys; Gln; Asn	Lys
Asn (N)	Gln; His; Lys; Arg	Gln
Asp (D)	Glu	Glu
Cys (C)	Ser	Ser
Gln (Q)	Asn	Asn
Glu (E)	Asp	Asp
Gly (G)	Pro; Ala	Ala
His (H)	Asn; Gln; Lys; Arg	Arg
Ile (I)	Leu; Val; Met; Ala; Phe	Leu
Leu (L)	Ile; Val; Met; Ala; Phe	Ile
Lys (K)	Arg; Gln; Asn	Arg
Met (M)	Leu; Phe; Ile	Leu
Phe (F)	Leu; Val; Ile; Ala; Tyr	Leu
Pro (P)	Ala	Ala
Ser (S)	Thr	Thr
Thr (T)	Ser	Ser
Trp (W)	Tyr; Phe	Tyr
Tyr (Y)	Trp; Phe; Thr; Ser	Phe
Val (V)	Ile; Leu; Met; Phe; Ala	Leu

10 本发明还包括 HDGF5 多肽编码序列及其片段的反义序列。这种反义序列可用于抑制细胞内 HDGF5 的表达。

本发明还包括可用作探针和引物的核酸分子, 该分子通常具有 HDGF5 多肽编码序列的约 8-100 个, 较佳地约 15-50 个连续核苷酸。该探针可用于检测样品中是否存在编码 HDGF5 的核酸分子。

本发明还包括检测 HDGF5 核苷酸序列的方法，它包括用上述的探针与样品进行杂交，然后检测探针是否发生了结合。较佳地，该样品是 PCR 扩增后的产物，其中 PCR 扩增引物对应于 HDGF5 多肽的编码序列，并可位于该编码序列的两侧或中间。引物长度一般为 20-50 个核苷酸。

5 在本发明中，可选用本领域已知的各种载体，如市售的载体。比如，选用市售的载体，然后将编码本发明多肽的核苷酸序列可操作地连于表达调控序列，可以形成蛋白表达载体。

如本文所用，“可操作地连于”指这样一种状况，即线性 DNA 序列的某些部分能够影响同一线性 DNA 序列其他部分的活性。例如，如果信号肽 DNA 作为前体表达并参与多肽的分泌，那么信号肽(分泌前导序列)DNA 就是可操作地连于多肽 DNA；如果启动子控制序列的转录，那么它是可操作地连于编码序列；如果核糖体结合位点被置于能使其翻译的位置时，那么它是可操作地连于编码序列。一般，“可操作地连于”意味着相邻近，而对于分泌前导序列则意味着在阅读框中相邻。

15 在本发明中，术语“宿主细胞”包括原核细胞和真核细胞。常用的原核宿主细胞的例子包括大肠杆菌、枯草杆菌等。常用的真核宿主细胞包括酵母细胞，昆虫细胞、和哺乳动物细胞。较佳地，该宿主细胞是真核细胞，如 CHO 细胞、COS 细胞等。

本发明还包括对 HDGF5 DNA 或是其片段编码的多肽具有特异性的多克隆抗体和单克隆抗体，尤其是单克隆抗体。这里，“特异性”是指抗体能结合于 HDGF5 基因产物或片段。较佳地，指那些能与 HDGF5 基因产物或片段结合但不识别和结合于其它非相关抗原分子的抗体。本发明中抗体包括那些能够结合并抑制 HDGF5 蛋白的分子，也包括那些并不影响 HDGF5 蛋白功能的抗体。本发明还包括那些能与修饰或未经修饰形式的 HDGF5 基因产物结合的抗体。

25 本发明不仅包括完整的单克隆或多克隆抗体，而且还包括具有免疫活性的抗体片段，如 Fab'或(Fab)<sub>2</sub>片段；抗体重链；抗体轻链；遗传工程改造的单链 Fv 分子；或嵌合抗体，如具有鼠抗体结合特异性但仍保留来自人的抗体部分的抗体。

本发明的抗体可以通过本领域内技术人员已知的各种技术进行制备。例如，纯化的 HDGF5 基因产物或者其具有抗原性的片段，可被施用于动物以诱导多克隆抗体的产生。与之相似的，表达 HDGF5 或其具有抗原性的片段的细胞可用来免疫动物来生产抗体。本发明的抗体也可以是单克隆抗体。此类单克隆抗体可以利用杂交瘤技术来制备(见 Kohler 等人, Nature 256:495, 1975; Kohler 等人, Eur.J.Immunol. 6:511, 1976; Kohler 等人, Eur.J.Immunol. 6:292, 1976; Hammerling 等人, In Monoclonal Antibodies and T Cell Hybridomas, Elsevier, N.Y., 1981)。本发

明的抗体包括能阻断 HDGF5 功能的抗体以及不影响 HDGF5 功能的抗体。本发明的各类抗体可以利用 HDGF5 基因产物的片段或功能区，通过常规免疫技术获得。这些片段或功能区可以利用重组方法制备或利用多肽合成仪合成。与 HDGF5 基因产物的未修饰形式结合的抗体可以用原核细胞(例如 E. coli)中生产的基因产物来免疫动物而产生；与翻译后修饰形式结合的抗体(如糖基化或磷酸化的蛋白或多肽)，可以用真核细胞(例如酵母或昆虫细胞)中产生的基因产物来免疫动物而获得。

本发明的人HDGF5核苷酸全长序列或其片段通常可以用PCR扩增法、重组法或人工合成的方法获得。对于PCR扩增法，可根据本发明所公开的有关核苷酸序列，尤其是开放阅读框序列来设计引物，并用市售的cDNA库或按本领域技术人员已知的常规方法所制备的cDNA库作为模板，扩增而得有关序列。当序列较长时，常常需要进行两次或多次PCR扩增，然后再将各次扩增出的片段按正确次序拼接在一起。

一旦获得了有关的序列，就可以用重组法来大批量地获得有关序列。这通常是将其克隆入载体，再转入细胞，然后通过常规方法从增殖后的宿主细胞中分离得到有关序列。

此外，还可用人工合成的方法来合成有关序列，尤其是片段长度较短时。通常，通过先合成多个小片段，然后再进行连接可获得序列很长的片段。

目前，已经可以完全通过化学合成来编码本发明蛋白(或其片段，或其衍生物)的 DNA 序列。然后可将该 DNA 序列引入本领域中的各种 DNA 分子(如载体)和细胞中。此外，还可通过化学合成将突变引入本发明蛋白序列中。

本发明蛋白的片段除了可用重组法产生之外，还可利用固相技术通过直接合成肽而加以生产(Stewart 等人, (1969) Solid-Phase Peptide Synthesis, WH Freeman Co., San Francisco; Merrifield J. (1963) J. Am Chem. Soc 85: 2149-2154)。在体外合成蛋白质可以用手工或自动进行。例如，可以用 Applied Biosystems 的 431A 型肽合成仪(Foster City, CA)来自动合成肽。可以分别化学合成本发明蛋白的各片段，然后用化学方法加以连接以产生全长的分子。

本发明蛋白的编码序列还可用于基因定位。例如，通过荧光原位杂交技术(FISH)，将 cDNA 克隆与分裂中期的染色体进行杂交，可以准确地进行染色体定位。该技术可以使用短至约 500bp 的 cDNA；也可以使用长至约 2000bp 或者更长的 cDNA。对于该技术，可参见 Verma 等人, Human Chromosomes: A Manual of Basic Techniques, Pergamon Press, New York(1988)。

一旦序列被定位于染色体上的某个精确位置，将可以将序列在染色体上的物理位置与遗传图谱数据相关联。这些遗传图谱数据是可以获得的，例如通过孟德



尔(Mendelian)人遗传数据库(可通过 Johns Hopkins University Welch Medical Library 在网上获得)。然后,通过连锁分析来鉴定基因与已定位于同一染色体区域的疾病之间的相关性。

接着,有必要确定患病个体和健康个体之间的 cDNA 或基因组序列方面的差异。如果某一突变存在于部分或全部患病个体但不存在于正常个体,那么该突变可能就是该疾病的致病因素。

利用本发明蛋白,通过各种常规筛选方法,可筛选出与 HDGF5 发生相互作用的物质,如受体、抑制剂或拮抗剂等。

本发明蛋白及其抗体、抑制剂、拮抗剂或受体等,当在治疗上进行施用(给药)时,可提供不同的效果。通常,可将这些物质配制于无毒的、惰性的和药学上可接受的水性载体介质中,其中 pH 通常约为 5-8,较佳地 pH 约为 6-8,尽管 pH 值可随被配制物质的性质以及待治疗的病症而有所变化。配制好的药物组合物可以通过常规途径进行给药,其中包括(但不限于):肌肉、腹膜内、皮下、皮内、或局部给药。

以本发明的人 HDGF5 蛋白为例,可以将其与合适的药学上可接受的载体联用。这类药物组合物含有治疗有效量的蛋白质和药学上可接受的载体或赋形剂。这类载体包括(但不限于):盐水、缓冲液、葡萄糖、水、甘油、乙醇、及其组合。药物制剂应与给药方式相匹配。本发明的人 HDGF5 蛋白可以被制成针剂形式,例如用生理盐水或含有葡萄糖和其他辅剂的水溶液通过常规方法进行制备。诸如片剂和胶囊之类的药物组合物,可通过常规方法进行制备。药物组合物如针剂、溶液、片剂和胶囊宜在无菌条件下制造。活性成分的给药量是治疗有效量,例如每天约 1 微克/千克体重-约 10 毫克/千克体重。此外,本发明的多肽还可与其他治疗剂一起使用。

当本发明的人 HDGF5 蛋白多肽被用作药物时,可将治疗有效剂量的该多肽施用于哺乳动物,其中该治疗有效剂量通常至少约 1 微克/千克体重,而且在大多数情况下不超过约 10 毫克/千克体重,较佳地该剂量是约 10 微克/千克体重-约 1 毫克/千克体重。当然,具体剂量还应考虑给药途径、病人健康状况等因素,这些都是熟练医师技能范围内的。

在本发明的一个具体实施方案中,人 HDGF5 的 cDNA 核苷酸序列是如此获得的,以人睾丸 λ gt11cDNA 文库(购自 Clontech 公司)为模板,用寡核苷酸 A1: 5'-cgctatgtcttgcttcagccg-3'(SEQ ID NO:3)为正向引物,寡核苷酸 A2: 5'-ggccctagcgggtttcccaag-3'(SEQ ID NO:4)为反向引物,进行 PCR,PCR 条件均为 93 °C 4 分钟,随之以 93 °C 1 分钟,68 °C 1 分钟和 72 °C 1 分钟进行 35 个循环,最

后 72 °C 延伸 5 分钟。电泳检测得到的 PCR 片段，约为 1000 个碱基的长度，为所需的目的片段。本发明的分离的多核苷酸全长为 990 个核苷酸，其详细序列见 SEQ ID NO:1，其中开放读框位于 5-910 位核苷酸。

5 肝癌细胞衍生生长因子(HDGF)是从人的肝癌细胞株 HuH-7 中分离到的肝素结合蛋白,它具有刺激细胞生长的活性,能够促进成纤维细胞和一些肝癌细胞的生长, HDGF 的有丝分裂原活性使其在急性恶性肝炎和肝损伤的治疗上存在着极大的应用价值。

研究表明, HDGF 虽然源自肝癌细胞株,但它在正常细胞和癌细胞系中均有广泛的表达,因而 HDGF 不仅仅局限于作用癌细胞,可能对正常的组织和细胞的生理和发育过程也有重要的生理作用。如:可促进 SMC、猴 Cos-7、Swiss-3T3、HuH-7、7.1.1 细胞分裂;可以促进与调节血管内皮细胞和血管平滑肌细胞的生长,可作为分泌性的生长因子或核因子发挥功能。HDGF 家族的已知各成员的表达模式是不同的,但是它们在睾丸中都有很高程度的富集,而且它们的 5'非翻译区均有高于 70%的 GC 比,因而在雄性生殖细胞发育过程中有重要功能,并和 DNA 甲基化,染色质构象以及翻译调控相关。

因此,本发明的 HDGF5 在肝组织部分切除后有助于促进肝细胞的增殖,加速创面愈合;通过 HDGF5 寻找其小分子拮抗物或抑制剂并制成药物,可对肝癌细胞起到抑制生长及诱导凋亡的作用,从而减缓并控制肝癌病情的发展。由于 HDGF5 在不同的细胞中表达的情况不同,故将 HDGF5 蛋白、其编码核酸或其活性部分制成试剂盒,可用于良、恶性肿瘤的早期诊断,还可对肝癌的术后疗效和预后监测起到辅助作用。

本发明的 HDGF5 对各种细胞都有不同程度的有丝分裂作用,因此还可用于调节血管内皮细胞和血管平滑肌生长,在促进创伤愈合、治疗器官的急性炎症性损伤、治疗肝癌以外的恶性肿瘤方面也有一定的应用价值。

此外,由于本发明的人 HDGF5 具有源自人的天然氨基酸序列,因此,与来源于其他物种的同族蛋白相比,预计在施用于人时将具有更高的活性和/或更低的副作用(例如在人体内的免疫原性更低或没有)。

30 下面结合具体实施例,进一步阐述本发明。应理解,这些实施例仅用于说明本发明而不适用于限制本发明的范围。下列实施例中未注明具体条件的实验方法,通常按照常规条件如 Sambrook 等人,分子克隆:实验室手册(New York: Cold Spring Harbor Laboratory Press, 1989)中所述的条件,或按制造厂商建议的条件。

### 实施例 1

## HDGF5 的 cDNA 序列的克隆和测定

### 1. 引物扩增

以人睾丸  $\lambda$  gt11cDNA 文库(购自 Clontech 公司)为模板, 用寡核苷酸 A1 : 5'- cgctatgtcttgccttcagccg -3'(SEQ ID NO:3)为正向引物, 寡核苷酸 A2 : 5'- ggccctagcgggtttcccaag -3'(SEQ ID NO:4)为反向引物, 进行 PCR, PCR 条件均为 93 °C 4 分钟, 随之以 93 °C 1 分钟, 68 °C 1 分钟和 72 °C 1 分钟进行 35 个循环, 最后 72 °C 延伸 5 分钟。电泳检测得到的 PCR 片段, 约为 1000 个碱基的长度, 即所需的目 的片段。

### 2. PCR 产物的测序

10 将如上获得的目的 cDNA 序列与 pGEM-T<sup>TM</sup> 载体(Promega)连接, 转化大肠杆菌 JM103, 用 QIAprep Plasmid 试剂盒(QIAGEN)提取质粒, 用双链嵌套式缺失试剂盒(Pharmacia)对插入片段进行定向系列缺失, 然后用 PCR 对缺失子进行快速鉴定及排序。用 SequiTherm EXCEL<sup>TM</sup> DNA 测序试剂盒(Epicentre Technologies)对依次截短的缺失子进行测序, 获得全长 cDNA 序列, 共 990bp, 详细序列见 SEQ

15 ID NO:1。

```

cgctatgtct tgccttcagcc gcccaaaata caagaccggg gacctggtgt ttgccaaatt      60
aaagggtctat gccattggc cagcgaggat tgaacatgtc actgaaccca accgctacca      120
ggtgttcttc ttcgggaccc atgagaccgc cctgctgggc cccaagcacc tttttcctta      180
tgaggagtcc aaggagaggt tcggcaagcc taacaagagg cgcggcttca gtgaggggct      240
gtgggagatc gagcacgacc ctatggctga ggcctcccct tgctgtgcc cagatgagga      300
gcagctttgt gccgaggagc cagggccagg agaggagcca gagccggggc aggagctgga      360
gccggaatcc aggctgagc tggaaatccat gcctgagctg gaggcagaac cgaggcctga      420
gaaagagtgt gagcaggagc cggagcagga gccgggagcag gagctggagc aggagccgga      480
gctggagccg gagccggagc cggagccgga gccgggagccg gagcccgagc ccgagccgga      540
gccggagccc cagcctgcct atgacctact ggatgccaag gaggagcctg gcctcattga      600
ggccgagcca ggagatcagc aagccgagca agtgcgagag cagcacgctg aagctgaggt      660
catggctgta gtggaggagc cggagagtct gaagaggagc gcggaggatg aacagcctca      720
cagtccctcc aaacggccca gggaggcggc gcctggcgcg ctggagatgg agccggctgg      780
agagcgcgag gcagaggcct gcccttcgt ggaggagcct gaccaagccc aggaacagca      840
gactccgttg gaagaagagg ccacagagga ggcagtccag ggcctgatgg ttggagaaat      900
cgaaggcctg tagtcacggt gtctgtaaaa gagccctctc taccggttcc tggtgccacc      960
tggtgtggc ttgggaaacc cgctagggcc      990

```

(SEQ ID NO: 1)

35 根据得到的全长 cDNA 序列推导出 HDGF5 的氨基酸序列, 共 302 个氨基酸残基, 其氨基酸序列详见 SEQ ID NO:2。

```

MSCFSRPKYK TGDLVFAKLK GYAHWPARI EHVTEPNRYQV FFFGTHETAL 50
LGPKHLFPYE ESKERFGKPN KRRGFSEGLW EIEHDPMAEA SPCLCPDEEQ 100
LCAEPPGPG EPEPGQLEP ESRPELESMP ELEAEPRPEK ECEQEPEQEP 150
EQUEQEPEL EPEPEPEPEP EPEPEPEPEP EPQPAYDLLD AKEEPGLIEA 200
EPGDQQAQV REQHAEEVM AVVEEPESLK RSAEDEQPHS PPKRPREAAP 250

```

GALEMEPAGE REAEACPFVE EPDQAQEQQT PLEEEATEEA VQGLMVGEIE 300  
GL 302

(SEQ ID NO:2)

## 5 实施例 2

### 同源比较

用人 HDGF5 的编码蛋白在 Non-redundant GenBank+ EMBL +DDBJ+PDB 数据库及 Non-redundant GenBank CDS translations+PDB+SwissProt+Spupdate+PIR 数据库中用 BLAST 进行蛋白同源检索。HDGF 家族成员氨基末端(HATH region ,  
10 for homologous to the amino terminus of the HDGF)具有高度保守性,被认为是该家族成员的特征性序列,而本发明的人 HDGF5 其氨基末端与 HDGF 家族其他成员都有较高同源性,参见附图中人 HDGF5 与 HDGF 家族其他成员氨基酸 N 端序列同源比较。鉴于本发明的上述特征认为其属于肝癌细胞衍生生长因子(HDGF)家族,具有与该家族成员相似的功能。

15 肝癌细胞衍生生长因子(HDGF)是从人的肝癌细胞株 HuH-7 中分离到的肝素结合蛋白,它具有刺激细胞生长的活性,能够促进成纤维细胞和一些肝癌细胞的生长, HDGF 的有丝分裂原活性使其在急性恶性肝炎和肝损伤的治疗上存在着极大的应用价值。

研究表明, HDGF 虽然源自肝癌细胞株,但它在正常细胞和癌细胞系中均有  
20 广泛的表达,因而 HDGF 不仅仅局限于作用癌细胞,可能对正常的组织细胞也有重要的生理作用。如:可促进 SMC、猴 Cos-7、Swiss-3T3、HuH-7、7.1.1 细胞分裂;可以促进血管生成;可作为分泌性的生长因子或核因子发挥功能。

本发明的 HDGF5 还可用于与其他蛋白一起产生融合蛋白,比如与免疫球蛋白一起产生融合蛋白。此外,本发明 HDGF5 还可以与该家族的其他成员进行融  
25 合或交换片段,以产生新的蛋白,如将本发明 HDGF5 的 N 端与人的其它 HDGF 及鼠 HDGF 的 N 端进行交换,以产生新的活性更高或具有新特性的蛋白。

针对本发明 HDGF5 的抗体,用于筛选该家族的其他成员,或者用于亲和纯化相关蛋白(如该家族的其他成员)。

此外,本发明人 HDGF5 核酸(编码序列或反义序列)可以被引入细胞,以提  
30 高人 HDGF5 的表达水平或者抑制人 HDGF5 的过度表达。本发明的人 HDGF5 蛋白或其活性多肽片段可以施用于病人,以治疗或减轻因人 HDGF5 缺失、无功能或异常而导致的有关病症。此外,还可以用基于本发明的核酸序列或抗体进行有关的诊断或预后判断。

### 实施例 3

#### HDGF5 在大肠杆菌中的表达

在该实施例中，将编码 HDGF5 的 cDNA 序列用对应于该 DNA 序列的 5'和 3'-端的 PCR 寡核苷酸引物进行扩增。

5 PCR 反应中使用的 5'寡核苷酸引物序列为：

E1: 5'-gcaggatccatgtcttgcttcagccgc-3' (SEQ ID NO:5)

该引物含有 BamHI 限制性内切酶的酶切位点，在该酶切位点之后是由起始密码子开始的 HDGF5 编码序列的部分核苷酸序列；

3'端引物序列为：E2: 5'- ccgaagcttcaggccttcgatttctcc-3' (SEQ ID NO:6)

10 该引物含有 HindIII 限制性内切酶的酶切位点、翻译终止子和 HDGF5 的部分编码序列。

引物上的限制性内切酶 BamHI 、 HindIII 的酶切位点对应于细菌表达载体 pQE-9(Qiagen Inc., Chatsworth, CA)上的限制性内切酶酶切位点，该质粒载体编码抗生素抗性(Amp<sup>r</sup>)、一个细菌复制起点(ori)、一个 IPTG-可调启动子/操纵子(P/O)、  
15 一个核糖体结合位点(RBS)、一个 6-组氨酸标记物(6-His)以及限制性内切酶克隆位点。

用 BamHI 和 HindIII 消化 pQE-9 载体及插入片段混合物，随后进行连接。再用连接混合物转化购自 Qiagen，商品名为 M15/rep4 的 E.coli 菌株，M15/rep4 含有多拷贝的质粒 pREP4，其表达 lacI 阻遏物并携带卡那霉素抗性(Kan<sup>r</sup>)。在含有  
20 Amp 和 Kan 的 LB 培养皿上筛选转化子，抽提质粒，用 EcoNI 酶切鉴定插入片段大小及方向，并测序验证 HDGF5 的 cDNA 片段已正确插入了载体。

在补加 Amp(100  $\mu$ g/ml)和 Kan(25  $\mu$ g/ml)的 LB 液体培养基中过夜培养(O/N)含所需构建物的阳性转化子克隆。过夜(O/N)培养物以 1 : 100-1 : 250 的稀释率稀释，然后接种到大体积培养基中，培养细胞生长至 600 光密度(OD<sub>600</sub>)为 0.4-0.6  
25 时，加入 IPTG(“异丙基硫代- $\beta$ -D-半乳糖苷”)至终浓度为 1mM。通过使 lacI 阻遏物失活，IPTG 诱导启动 P/O 导致基因表达水平提高。继续培养细胞 3-4 小时，随后离心(6000  $\times$  g，20 分钟)。超声裂解培养物，收集细胞裂解液并将其稀释于 6M 盐酸胍中。澄清后，通过在能使含 6-His 标记物蛋白紧密结合的条件下，用镍-螯合柱层析从溶液中纯化溶解的 HDGF5。用 6M 盐酸胍(pH5.0)从柱中洗脱  
30 HDGF5。可用几种方法从盐酸胍中变性沉淀蛋白。或者使用透析步骤除去盐酸胍，或者从镍-螯合柱中分离出纯化蛋白，纯化后的蛋白可以结合到第二个柱中，该柱中具有递减的线性盐酸胍梯度。在结合到该柱时蛋白质变性，随后用盐酸胍(pH5.0)洗脱。最后，将可溶的蛋白质对含 PBS 进行透析，然后将蛋白质保存在终浓度为 10%(w/v)甘油的贮存液中。

用 12% 的 SDS-PAGE 胶进行电泳, 鉴定表达蛋白的分子量大小约为 34148。

此外, 用常规方法对表达蛋白的 N 端和 C 端各 10 个氨基酸长度的氨基酸进行测序, 发现与 SEQ ID NO:2 的序列一致。

#### 5 实施例 4

##### HDGF5 在真核细胞(CHO 细胞株)中的表达

在该实施例中, 将编码 HDGF5 的 cDNA 序列用对应于该 DNA 序列的 5' 和 3' 端的 PCR 寡核苷酸引物进行扩增, 获得 HDGF5 cDNA 作为插入片段。

PCR 反应中使用的 5' 寡核苷酸引物序列为:

10 C1: 5'- gcaaagcttatgtcttgcttcagccgc-3' (SEQ ID NO: 7)

该引物含有 HindIII 限制性内切酶的酶切位点, 在该酶切位点之后是由起始密码子开始的 HDGF5 编码序列的 18 个核苷酸;

3' 端引物序列为:

C2: 5'- ccgggatcccaggccttcgattctcc -3' (SEQ ID NO: 8)

15 该引物含有 BamHI 限制性内切酶的酶切位点、一个翻译终止子和 HDGF5 的部分编码序列。

引物上的限制性内切酶的酶切位点对应于 CHO 细胞表达载体 pcDNA3 上的限制性内切酶酶切位点, 该质粒载体编码抗生素抗性(Amp<sup>r</sup> 和 Neo<sup>r</sup>)、一个噬菌体复制起点(f1 ori)、一个病毒复制起点(SV40 ori)、一个 T7 启动子、一个病毒启动子(P-CMV)、一个 Sp6 启动子、一个 SV40 启动子、一个 SV40 加尾信号和相应的  
20 polyA 顺序、一个 BGH 加尾信号和相应的 polyA 顺序。

用 HindIII 和 BamHI 消化 pcDNA3 载体及插入片段, 随后将插入片段连接到 pcDNA3 载体。随后用连接混合物转化 E.coli DH5  $\alpha$  菌株。在含有 Amp 的 LB 培养皿上筛选转化子, 在补加 Amp(100  $\mu$  g/ml)的 LB 液体培养基中过夜培养(O/N)  
25 含所需构建物的克隆。抽提质粒, 用 EcoNI 酶切鉴定插入片段大小及方向, 并测序验证 HDGF5 的 cDNA 片段已正确插入了载体。

质粒转染 CHO 细胞是用脂转染法, 用 Lipofectin 试剂盒(GiBco Life)进行。转染 48 小时后, 经 2-3 周的持续 G418 加压筛选, 收集细胞及细胞上清测定表达蛋白酶活力。去 G418, 连续传代培养; 对混合克隆细胞极限稀释, 选择具有较  
30 高蛋白活性的细胞亚克隆。按常规方法大量培养上述阳性亚克隆。48 小时后, 开始收集细胞及上清, 用超声裂解方法破碎细胞。以含 0.05% Triton 的 50mM Tris-HCl(pH7.6)溶液为平衡液及洗脱液, 用经预平衡的 Superdex G-75 柱收集上述蛋白的活性峰。再用 50mM Tris-HCl(pH8.0)平衡的 DEAE-Sepharose 柱, 以含 0-1M NaCl 的 50mM Tris-HCl(pH8.0)溶液为洗脱液进行梯度洗脱, 收集上述蛋

白的活性峰。然后以 PBS(pH7.4)为透析液对表达蛋白溶液进行透析。最后冻干保存。

用 12%的 SDS-PAGE 胶进行电泳, 鉴定表达蛋白的分子量大小为 34148。

此外, 用常规方法对表达蛋白的 N 端和 C 端各 10 个氨基酸长度的氨基酸进行测序, 发现与 SEQ ID NO:2 的序列一致。

## 实施例 5

### 制备抗体

将实施例 3 和 4 获得的重组蛋白用来免疫动物以产生抗体, 具体如下。重组分子用层析法进行分离后备用。也可用 SDS-PAGE 凝胶电泳法进行分离, 将电泳条带从凝胶中切下, 并用等体积的完全 Freund's 佐剂乳化。用 50-100  $\mu$ g/0.2ml 乳化过的蛋白, 对小鼠进行腹膜内注射。14 天后, 用非完全 Freund's 佐剂乳化的同样抗原, 对小鼠以 50-100  $\mu$ g/0.2ml 的剂量进行腹膜内注射以加强免疫。每隔 14 天进行一次加强免疫, 至少进行三次。获得的抗血清的特异反应活性用它在体外沉淀 HDGF5 基因翻译产物的能力加以评估。结果发现, 抗体可特异性地与本发明蛋白发生沉淀。

## 实施例 6

### HDGF5 对细胞增殖的促进作用

通过在细胞培养液中加入不同浓度的 HDGF5 的方法检测 HDGF5 对细胞增殖的作用。具体步骤如下: 用 10 % 1640 培养液将小鼠内皮细胞 Berd-3 细胞接种于 96 孔板, 使细胞浓度达到  $1-1.2 \times 10^4$  个 / ml; 将细胞放置在 37  $^{\circ}$ C 培养箱 (5 %  $\text{CO}_2$ ) 中过夜, 然后换无血清培养基于 37  $^{\circ}$ C, 5 %  $\text{CO}_2$  条件下再过一夜; 用纯化的实施例 4 中制得的 HDGF5 蛋白, 加入上述培养好的小鼠内皮细胞培养孔中, 使各孔内 HDGF5 的浓度呈梯度排列, 即 0.0 (空白对照)、0.1ng/ml、0.5ng/ml、2.5ng/ml、25ng/ml、100ng/ml。上述加入了 HDGF5 的细胞再于 37  $^{\circ}$ C, 5 %  $\text{CO}_2$  条件下培养, 并分别于 24、48、72 小时停止培养, 取出进行 MTS 检测。

检测结果显示, 随着 HDGF5 含量的增加, 细胞数量也相应增多。

在本发明提及的所有文献都在本申请中引用作为参考, 就如同每一篇文献被单独引用作为参考那样。此外应理解, 在阅读了本发明的上述讲授内容之后, 本领域技术人员可以对本发明作各种改动或修改, 这些等价形式同样落于本申请所附权利要求书所限定的范围。

## 权 利 要 求 书

1. 一种分离出的DNA分子，其特征在于，它包括：编码具有人HDGF5蛋白活性的多肽的核苷酸序列，所述的核苷酸序列与SEQ ID NO:1中从核苷酸5-910位的核苷酸序列有至少70%的同源性；或者所述的核苷酸序列能在中度严紧条件下与SEQ ID NO:1中从核苷酸5-910位的核苷酸序列杂交。
- 2.如权利要求1所述的DNA分子，其特征在于，所述的序列编码一多肽，该多肽具有SEQ ID NO:2所示的序列。
- 3.如权利要求1所述的DNA分子，其特征在于，该序列具有SEQ ID NO:1中核苷酸5-910位的序列。
- 4.一种分离的HDGF5蛋白多肽，其特征在于，它包括：具有SEQ ID NO:2氨基酸序列的多肽、或其保守性变异多肽、或其活性片段、或其活性衍生物。
- 5.如权利要求4所述的多肽，其特征在于，该多肽是具有SEQ ID NO:2序列的多肽。
- 6.一种载体，其特征在于，它含有权利要求1所述的DNA。
- 7.一种宿主细胞，其特征在于，它被权利要求6所述载体转化。
- 8.一种产生具有HDGF5蛋白活性的多肽的方法，其特征在于，该方法包括：
  - (a)将编码具有HDGF5蛋白活性的多肽的核苷酸序列可操作地连于表达调控序列，形成HDGF5蛋白表达载体，所述的核苷酸序列与SEQ ID NO:1中从核苷酸5-910位的核苷酸序列有至少70%的同源性；
  - (b)将步骤(a)中的表达载体转入宿主细胞，形成HDGF5蛋白的重组细胞；
  - (c)在适合表达HDGF5蛋白多肽的条件下，培养步骤(b)中的重组细胞；
  - (d)分离出具有HDGF5蛋白活性的多肽。
- 9.一种抗体，其特征在于，它与HDGF5蛋白多肽特异性结合。
- 10.一种探针分子，其特征在于，它含有具有如 SEQ ID NO:1 所示的 DNA 分子中 8-100 个连续核苷酸。



人 HDGF5 MSCFSRPK-YKTGDLVFAKLKGYAHWPARIEHVTEP-----NRYQVFFFGTHETAFLGP 53  
 大鼠 HRP MSCFSRPK-YKTGDLVFAKLKGYAHWPARIEHVTEP-----NRYQVFFFGTHETALLGP 53  
 小鼠 HRP1 MSCFSRSK-YKTGDLVFAKLKGYAHWPARIEHVAEA-----NRYQVFFFGTHETALLGP 53  
 小鼠 HDGF MSRSNRQKEYKCGDLVFAKMKGYPHWPARIDEMPEAAVKSTANKYQVFFFGTHETAFLGP 60  
 大鼠 HDGF MSRSNRQKEYKCGDLVFAKMKGYPHWPARIDEMPEAAVKSTANKYQVFFFGTHETAFLGP 60  
 人 HDGF MSRSNRQKEYKCGDLVFAKMKGYPHWPARIDEMPEAAVKSTANKYQVFFFGTHETAFLGP 60  
 人 HRP1 MSRSNRQKEYKCGDLVFAKMKGYPHWPARIDEMPEAAVKSTANKYQVFFFGTHETAFLGP 60  
 人 HRP3 -MARPRPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFFGTHETAFLGP 59  
 小鼠 HRP3 -MARPRPREYKAGDLVFAKMKGYPHWPARIDELPEGAVKPPANKYPIFFFGTHETAFLGP 59  
 小鼠 HRP2 -----MPHAFKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPNKYPIFFFGTHETAFLGP 55  
 大鼠 HDGF3 -----MPHAFKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPNKYPIFFFGTHETAFLGP 55

: :\* \*\*\*\*\*:\*\*\*.\*\*\*\*\*:. . . :                   \*:\* :\*\*\*\*\*:\*\*\*

人 HDGF5 KHLFPYEESKERFGKPNKRRGFSEGLWEIEHDPMAEASPCCLPDEEQLCAEEP GPGE EPE 113  
 大鼠 HRP KHLFPYEESKERFGKPNKRRGFSEGLWEIEHDP MVEASPCCLPDEEQLCAEEP GPGE EPE 113  
 小鼠 HRP1 RHLFPYEESKEKFGKPNKRRGFSEGLWEIEHDP MVEASSSLCSEEDQSYTEDPGLAEEPE 113  
 小鼠 HDGF KDLFPYEESKEKFGKPNKRRGFSEGLWEIENNP TVKASGYQSS-QKKSCAAEP----- 112  
 大鼠 HDGF KDLFPYEESKEKFGKPNKRRGFSEGLWEIENNP TVKASGYQSS-QKKSCAAEP----- 112  
 人 HDGF KDLFPYEESKEKFGKPNKRRGFSEGLWEIENNP TVKASGYQSS-QKKSCVEEP----- 112  
 人 HRP1 KDLFPYEESKEKFGKPNKRRGFSEGLWEIENNP TVKASGYQSS-QKKSCVEEP----- 112  
 人 HRP3 KDLFPYKEYKDKFGKSNKRRGFNEGLWEIENNP GVKFTGYQAIQQSSSETEGEGGNTAD 119  
 小鼠 HRP3 KDLFPYKEYKDKFGKSNKRRGFNEGLWEIENNP GVKFTGYQTIQQSSSETEGEGGNTAD 119  
 小鼠 HRP2 KDLFPYDKCKDKYGKPNKRRGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEADLGCGSD 115  
 大鼠 HDGF3 KDLFPYDKCKDKYGKPNKRRGFNEGLWEIQNNPHASYSAPLPVSSSDSEAPEADLGGGSD 115

..\*\*\*\*. : \*:::\*.\*\*\*:\*.\*\*\*\*\*:::\* . . :                   ... :

图 1

## 序列表

&lt;110&gt; 余, 龙

&lt;120&gt; 人类肝癌衍生生长因子 5、其编码序列、制法及用途

&lt;130&gt; 017216

&lt;150&gt; CN02110535.9

&lt;151&gt; 2002-01-11

&lt;160&gt; 8

&lt;170&gt; PatentIn version 3.1

&lt;210&gt; 1

&lt;211&gt; 990

&lt;212&gt; DNA

&lt;213&gt; 智人(Homo sapiens)

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (5)..(910)

&lt;223&gt;

&lt;400&gt; 1

cgct atg tct tgc ttc agc cgc cca aaa tac aag acc ggg gac ctg gtg	49
Met Ser Cys Phe Ser Arg Pro Lys Tyr Thr Gly Asp Leu Val	
1 5 10 15	

ttt gcc aaa tta aag ggc tat gcc cat tgg cca gcg agg att gaa cat	97
Phe Ala Lys Leu Lys Gly Tyr Ala His Trp Pro Ala Arg Ile Glu His	
20 25 30	

gtc act gaa ccc aac cgc tac cag gtg ttc ttc ttc ggg acc cat gag	145
Val Thr Glu Pro Asn Arg Tyr Gln Val Phe Phe Phe Gly Thr His Glu	
35 40 45	

acc gcc ctg ctg ggc ccc aag cac ctt ttt cct tat gag gag tcc aag	193
Thr Ala Leu Leu Gly Pro Lys His Leu Phe Pro Tyr Glu Glu Ser Lys	
50 55 60	

gag agg ttc ggc aag cct aac aag agg cgc ggc ttc agt gag ggg ctg	241
Glu Arg Phe Gly Lys Pro Asn Lys Arg Arg Gly Phe Ser Glu Gly Leu	
65 70 75	

tgg gag atc gag cac gac cct atg gct gag gcc tcc cct tgc ctg tgc	289
Trp Glu Ile Glu His Asp Pro Met Ala Glu Ala Ser Pro Cys Leu Cys	
80 85 90 95	

cca gat gag gag cag ctt tgt gcc gag gag cca ggg cca gga gag gag	337
Pro Asp Glu Glu Gln Leu Cys Ala Glu Glu Pro Gly Pro Gly Glu Glu	
100 105 110	

cca gag ccg ggg cag gag ctg gag ccg gaa tcc agg cct gag ctg gaa	385
Pro Glu Pro Gly Gln Glu Leu Glu Pro Glu Ser Arg Pro Glu Leu Glu	

115	120	125	
tcc atg cct gag ctg gag gca gaa ccg agg cct gag aaa gag tgt gag			433
Ser Met Pro Glu Leu Glu Ala Glu Pro Arg Pro Glu Lys Glu Cys Glu			
130	135	140	
cag gag ccg gag cag gag ccg gag cag gag ctg gag cag gag ccg gag			481
Gln Glu Pro Glu Gln Glu Pro Glu Gln Glu Leu Glu Gln Glu Pro Glu			
145	150	155	
ctg gag ccg gag ccg gag ccg gag ccg gag ccg gag ccg gag ccc gag			529
Leu Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu			
160	165	170	175
ccc gag ccg gag ccg gag ccc cag cct gcc tat gac cta ctg gat gcc			577
Pro Glu Pro Glu Pro Glu Pro Gln Pro Ala Tyr Asp Leu Leu Asp Ala			
180	185	190	
aag gag gag cct ggc ctc att gag gcc gag cca gga gat cag caa gcc			625
Lys Glu Glu Pro Gly Leu Ile Glu Ala Glu Pro Gly Asp Gln Gln Ala			
195	200	205	
gag caa gtg cga gag cag cac gct gaa gct gag gtc atg gct gta gtg			673
Glu Gln Val Arg Glu Gln His Ala Glu Ala Glu Val Met Ala Val Val			
210	215	220	
gag gag ccg gag agt ctg aag agg agc gcg gag gat gaa cag cct cac			721
Glu Glu Pro Glu Ser Leu Lys Arg Ser Ala Glu Asp Glu Gln Pro His			
225	230	235	
agt cct ccc aaa cgg ccc agg gag gcg gcg cct ggc gcg ctg gag atg			769
Ser Pro Pro Lys Arg Pro Arg Glu Ala Ala Pro Gly Ala Leu Glu Met			
240	245	250	255
gag ccg gct gga gag cgc gag gca gag gcc tgc ccc ttc gtg gag gag			817
Glu Pro Ala Gly Glu Arg Glu Ala Glu Ala Cys Pro Phe Val Glu Glu			
260	265	270	
cct gac caa gcc cag gaa cag cag act ccg ttg gaa gaa gag gcc aca			865
Pro Asp Gln Ala Gln Glu Gln Gln Thr Pro Leu Glu Glu Glu Ala Thr			
275	280	285	
gag gag gca gtc cag ggc ctg atg gtt gga gaa atc gaa ggc ctg			910
Glu Glu Ala Val Gln Gly Leu Met Val Gly Glu Ile Glu Gly Leu			
290	295	300	
tagtcacggt gtctgtaaaa gagccctctc taccggtcc tgggtgccacc tggctgtggc			970
ttgggaaacc cgctagggcc			990

&lt;210&gt; 2

&lt;211&gt; 302

&lt;212&gt; PRT

&lt;213&gt; 智人(Homo sapiens)

&lt;400&gt; 2

Met Ser Cys Phe Ser Arg Pro Lys Tyr Lys Thr Gly Asp Leu Val Phe  
 1 5 10 15

Ala Lys Leu Lys Gly Tyr Ala His Trp Pro Ala Arg Ile Glu His Val  
 20 25 30

Thr Glu Pro Asn Arg Tyr Gln Val Phe Phe Phe Gly Thr His Glu Thr  
 35 40 45

Ala Leu Leu Gly Pro Lys His Leu Phe Pro Tyr Glu Glu Ser Lys Glu  
 50 55 60

Arg Phe Gly Lys Pro Asn Lys Arg Arg Gly Phe Ser Glu Gly Leu Trp  
 65 70 75 80

Glu Ile Glu His Asp Pro Met Ala Glu Ala Ser Pro Cys Leu Cys Pro  
 85 90 95

Asp Glu Glu Gln Leu Cys Ala Glu Glu Pro Gly Pro Gly Glu Glu Pro  
 100 105 110

Glu Pro Gly Gln Glu Leu Glu Pro Glu Ser Arg Pro Glu Leu Glu Ser  
 115 120 125

Met Pro Glu Leu Glu Ala Glu Pro Arg Pro Glu Lys Glu Cys Glu Gln  
 130 135 140

Glu Pro Glu Gln Glu Pro Glu Gln Glu Leu Glu Gln Glu Pro Glu Leu  
 145 150 155 160

Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro Glu Pro  
 165 170 175

Glu Pro Glu Pro Glu Pro Gln Pro Ala Tyr Asp Leu Leu Asp Ala Lys  
 180 185 190

Glu Glu Pro Gly Leu Ile Glu Ala Glu Pro Gly Asp Gln Gln Ala Glu  
 195 200 205

Gln Val Arg Glu Gln His Ala Glu Ala Glu Val Met Ala Val Val Glu  
 210 215 220

Glu Pro Glu Ser Leu Lys Arg Ser Ala Glu Asp Glu Gln Pro His Ser  
 225 230 235 240

Pro Pro Lys Arg Pro Arg Glu Ala Ala Pro Gly Ala Leu Glu Met Glu  
 245 250 255

Pro Ala Gly Glu Arg Glu Ala Glu Ala Cys Pro Phe Val Glu Glu Pro  
 260 265 270

Asp Gln Ala Gln Glu Gln Gln Thr Pro Leu Glu Glu Glu Ala Thr Glu  
 275 280 285

Glu Ala Val Gln Gly Leu Met Val Gly Glu Ile Glu Gly Leu  
 290 295 300

<210> 3  
 <211> 21  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <222> (1)..(21)  
 <223> 引物

<400> 3  
 cgctatgtct tgcttcagcc g

21

<210> 4  
 <211> 21  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature  
 <222> (1)..(21)  
 <223> 引物

<400> 4  
 ggccctagcg ggtttccaa g

21

<210> 5  
 <211> 27  
 <212> DNA  
 <213> 人工序列

<220>  
 <221> misc\_feature

<222> (1)..(27)

<223> 引物

<400> 5

gcaggatcca tgtcttgctt cagccgc

27

<210> 6

<211> 27

<212> DNA

<213> 人工序列

<220>

<221> misc\_feature

<222> (1)..(27)

<223> 引物

<400> 6

ccgaagcttc aggccttcga tttctcc

27

<210> 7

<211> 27

<212> DNA

<213> 人工序列

<220>

<221> misc\_feature

<222> (1)..(27)

<223> 引物

<400> 7

gcaaagctta tgtcttgctt cagccgc

27

<210> 8

<211> 27

<212> DNA

<213> 人工序列

<220>

<221> misc\_feature

<222> (1)..(27)

<223> 引物

<400> 8

ccgggatccc aggccttcga tttctcc

27

## INTERNATIONAL SEARCH REPORT

PCT/CN03/00003

## A. CLASSIFICATION OF SUBJECT MATTER

IPC<sup>7</sup> C12N15/12

According to International Patent Classification(IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched(classification system followed by classification symbols)

IPC<sup>7</sup> C12N15/12

Documentation searched other than minimum documentation to the extent that such documents are included in the field searched

Electronic data base consulted during the international search(name of data base and, where practicable, search terms used)

CNPAT, WPI

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant claim No.
X	WO, A2, 0037492 (LILLY & CO ELI et al) 29.June 2000, See the whole document	1-10
X	WO, A1, 9639485 (HUMAN GENOME SCI INC et al) 12.December 1996, See the whole document	1-10
X	US, A, 5235042 (CHILDRENS MEDICAL CENT) 10.August 1993 , See the whole document	1-10
X	WO, A, 8706239 (GEN HOSPITAL CORP) 22.October 1987 , See the whole document	1-10



Further documents are listed in the continuation of Box C.



See patent family annex.

\* Special categories of cited documents:

"A"document defining the general state of the art which is not considered to be of particular relevance

"E"earlier document but published on or after the international filing date

"L"document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason(as specified)

"O"document referring to an oral disclosure, use, exhibition or other means

"P"document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;"document member of the same patent family

Date of the actual completion of the international search

25 March 2003(25.03.03)

Date of mailing of the international search report

05 JUN 2003 (05.06.03)

Name and mailing address of the ISA/

The Chinese Patent Office  
6, Xitucheng Road, Haidian District,  
Beijing, 100088, China

Facsimile No. 86-10-62019451

Authorized officer

PAN, aiquan

Telephone No. 86-10-62093906



### Information on patent family members

Patent document cited in search report	Publication date	Patent family members	Publication date
WO-A2-0037492	29-06-00	AU-A-200020595	12-07-00
WO-A1-9639485	12-12-96	EP- A1-0833892	08-04-98
US-A- 5235042	10-08-93	None	
WO-A-8706239	22-10-87	EP-A-0241830	21-10-87



## 国际检索报告

国际申请号

PCT/CN03/00003

## A. 主题的分类

Int.Cl<sup>7</sup>: C12N15/12

按照国际专利分类表(IPC)或者同时按照国家分类和 IPC 两种分类

## B. 检索领域

检索的最低限度文献(标明分类体系和分类号)

Int.Cl<sup>7</sup>: C12N15/12

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称和, 如果实际可行的, 使用的检索词)

WPI,CPRS

## C. 相关文件

类 型*	引用文件, 必要时, 包括相关段落的说明	相关的权利要求编号
X	WO, A2, 0037492 (伊莱利利等) 29.6 月 2000, 参见全文	1-10
X	WO, A1, 9639485 (人基因组科学公司等) 12.12 月 1996, 参见全文	1-10
X	US, A, 5235042 (儿童医学中心) 10.8 月 1993, 参见全文	1-10
X	WO, A, 8706239 (基因医院公司) 22.10 月 1987, 参见全文	1-10

☐ 其余文件在 C 栏的续页中列出。☒ 见同族专利附件。

\* 引用文件的专用类型:

"A" 明确表示了一般现有技术、不认为是特别相关的文件  
 "E" 在先文件, 但是在国际申请日的同一日或之后公布的  
 "L" 对优先权要求可能产生怀疑或者用来确定另一篇引用文件的公布日期或其它特殊理由而引用的文件(如详细说明)

"O" 涉及口头公开、使用、展览或其他手段的文件

"P" 在国际申请日之前但迟于所要求的优先权日公布的文件

"T" 在国际申请日或优先权日之后公布的在后文件, 它与申请不相抵触, 但是引用它是为了理解构成发明基础的理论或原理  
 "X" 特别相关的文件; 当该文件被单独使用时, 要求保护的发明不能认为是新颖的或不能认为具有创造性

"Y" 特别相关的文件; 当该文件与其他一篇或多篇这类文件结合在一起, 这种结合对本领域技术人员是显而易见的, 要求保护的发明不能认为具有创造性

"&amp;" 同族专利成员的文件

国际检索实际完成的日期

25.3 月 2003(25.03.03)

国际检索报告邮寄日期

05. 6月 2003 (05.06.03)

国际检索单位名称和邮寄地址

中国专利局

中国北京市海淀区西土城路 6 号(100088)

传真号:

86-010-62019451

受权官员:

潘爱群

电话号码: 86-10-62093906



## PCT/CN03/00003

PCT/ISA/210 表(同族专利附件)(1992 年 7 月)